Niveau(x) d'enseignement: 7 Numéro EDUQ : 8270214
Centre d'intérêt: 64

Keller, Arnold.

Computer-assisted instruction in the teaching of language skills : final report, 1977-78 / Arnold Keller with additional materials by Anne Blott. -- [Saint-Laurent] : Vanier College, 1978.
vi, 57 f. ; 28 cm.

Bibliographie: f. 54-57.

Cote Dewey: 371.3'9445/K29c Cote LC: LB/1028.5

Descripteur(s): *Anglais (Langue) *Composition *Enseignement correctif *Enseignement par ordinateur *Enseignement programmé *Manuel programmé *Evaluation d'une méthode *Collégial (Niveau)

L'étude a pour but de concevoir, de produire, d'utiliser et d'évaluer du matériel didactique appliqué à la composition anglaise. Elle vise à découvrir si l'utilisation de l'enseignement assisté par ordinateur permet un apprentissage plus efficace des mécanismes de l'écriture que d'autres types d'enseignement programmé. L'expérience portait sur l'enseignement correctif de l'anglais à des étudiants du collège Vanier, campus de Snoudon.

L'étude a été réalisée auprès de 75 étudiants du collégial, choisis au moyen du "Missouri College English test". Les étudiants ont été répartis au hasard en cinq groupes et soumis en 1978 à des types d'enseignement différents: 1. livre traditionnel; 2. livre brouillé; 3. livre brouillé avec exercices; 4. enseignement assisté par ordinateur sans exercices; 5. enseignement assisté par ordinateur avec exercices. Les résultats ont été mesurés à l'aide de trois tests standardisés.

L'expérience ne permet pas de conclure que l'enseignement assisté par ordinateur assure un apprentissage plus efficace des mécanismes de l'écriture, les résultats des cinq groupes d'étudiants étant presque identiques. En ce qui concerne la composition, l'enseignement assisté par ordinateur semble favoriser une amélioration légèrement plus grande chez les étudiants.

Type de document: r
Région(s) administrative(s): 063
Financement: Bénéficiaire - Collège Vanier
Organisme(s) source(s) - MEQ - DGEC
Programme(s) de subvention - PROSIP
Auteur de l'analyse: jr

E D U Q 8270214

E D U Q 8270214

COMPUTER-ASSISTED INSTRUCTION IN THE TEACHING OF LANGUAGE SKILLS

FINAL REPORT

1977-78

Vanier College
Arnold Keller
With additional materials by
Anne Blott
June 1978

ABSTRACT

Lessons in the mechanics of English composition were developed and presented in five different modes to first-year remedial students at Vanier College, Snowdon Campus. These five modes were: 1) CAI with drill-and-practice; 2) CAI without drill-and-practice; 3) printed scrambled PI with drill-and-practice; 4) printed scrambled PI without drill-and-practice; and 5) printed linear. No significant differences were found in test scores, averages for all essays submitted, averages for the last three essays, and attitude. CAI was found to be superior, however, from a teacher-designer's point of view because of its relative ease of implementation, its cost advantages given the conditions of the research, and its capacity to gather data on student performance.

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge a number of debts we incurred in the course of this year's work. Sarah Paltiel, Vanier College's Academic Dean, supported and encouraged the project from its very beginnings. Both Dr. Marc Scholer of SGME and Mr. Aaron Caplan of the University of Montreal gave generously of their time and advice. We especially wish to thank Dr. Gary Boyd of Concordia University for sharing with us his expertise in Educational Technology in general and Computer-Assisted Instruction in particular. We were also very fortunate to have an outstanding support staff led by Laura Winer whose diligence and care helped us greatly. And naturally, we wish to thank our students for their patience and cooperation. To all, our sincere thanks.

A.K.
A.B.
June 1978

TABLE OF CONTENTS

Abstract iii

Acknowledgements iv

Introduction 1

Problem Statement 5

The State of the Question 8

Objectives 17

Definitions 18

Population 19

Procedure 20

The Material 24

Design 32

Data and Instrumentation 33

Results 33

Discussion 38

Appendix I: Student Evaluation of Terminal and Booklet Grammar Lessons 49

Appendix II: Second-Language Students 51

References 54

LIST OF TABLES

Table I: Analysis of Variance of Test Scores for the Five Treatment Groups 33

Table II: Analysis of Variance Comparing Scores on Essay Averages for the Whole Semester 34

Table III: Analysis of Variance for Average Scores of Last Three Essays 34

Table IV: The Means and Standard Deviations of Test Scores for the Five Treatment Groups 34

Table V: Chi-square Analysis of Distribution of Subjects Without All Previous Education in English 35

Table VI: Percentage of Favourable Responses on Attitude Survey 36

Table VII: Analysis of Variance for Percentage of Favourable Responses on Attitude Survey 36

INTRODUCTION

The objective of this project has been to design, produce, implement, and evaluate a set of materials to teach the mechanics of English composition. To this end, a series of twelve lessons was developed and presented to 75 students enrolled in four remedial English courses at Vanier College. The primary research problem studied was which mode of delivering instruction to students resulted in highest achievement. Therefore, the materials were presented in five parallel modes: 1) a printed linear text; 2) a printed, scrambled text with drill-and-practice sequences (hereafter known as D&P); 3) a printed scrambled text without D&P; 4) a computer-assisted program with D&P sequences; and 5) a computer-assisted program without D&P. Design and production began in September 1977, and implementation and evaluation took place from February to May 1978. Computer lessons were in an extended version of Dartmouth Basic and installed on the CDC 6400 at Concordia University.

The question examined has particular significance for a publicly-financed institution like Vanier College which has the difficult task of providing the highest possible level of education to many students of various backgrounds. At the same time, the College must find the most efficient means of disbursing its limited human and financial resources. In the teaching of language skills, there is a particular need to upgrade abilities. Since it is clear that such skills are important to students

in the short term (to succeed in current studies) and in the long term (to become responsible and articulate members of the community), it is distressing to find that fewer and fewer entering students can write adequately. The situation is neither especially recent nor confined to Vanier College; and since other disciplines do not appear to insist that their students demonstrate strong communication skills, the task of improving writing generally has fallen to English departments--themselves lacking the resources to provide each student the personal supervision that we presume to be the best kind of training.

Computer-Assisted Instruction (hereafter known as CAI) and Programmed Instruction (hereafter known as PI) offer at least partial solutions. A programmed text--whether linear or branched--recognizes the need for individualized instruction and practice. A computer-assisted program theoretically is an even more powerful tool because it can deliver tutorial and D&P routines that respond more completely to personal styles and rates of learning. The limitations of both modes must be recognized clearly, however; given the state of the art in both CAI and PI, they simply cannot perform as well as human beings when it comes to the more complex tasks involved in teaching the higher cognitive skills. But within carefully defined areas, CAI and PI can perform quite well, effecting considerable savings in teacher time and dollars.

PI can be divided into linear and branching modes. The linear approach derives from the work of B.F. Skinner, and is based on the principles of operant conditioning. All students proceed through instruction in the same

sequence, the only variations being the rate at which they complete the material. Subject matter is broken down into relatively small increments (or frames) which are presented in the order which seems most logical to the designer. (Early designers of programmed texts followed principles enunciated by Skinner, and later N.A. Crowder; at present, most programmed texts appear quite eclectic in psychological orientation.) A PI program requires the learner to make frequent responses--usually one per frame--and programs are so written that relatively few student errors occur. This is accomplished through the use of prompts, a type of help which guides the student to the right answer without simply stating it directly. A linear program provides immediate knowledge of results since correct answers appear in the frame following the one in which a question is asked. Such knowledge reinforces the student's behavior, and theoretically, a program that elicits few errors will offer strong positive reinforcement.

In the branching mode (whether in a scrambled textbook or on a computer), students vary not only the rate at which they complete instruction, but also the order they go through it. This is because the branching sequences direct the student to different areas of instruction, depending upon the response made. Such instruction most often has been remedial tutoring (Rockert and Morton, p. 104), in which the same concept is taught using a strategy different than the initial presentation. Or a student may be branched to a series of D&P questions until he demonstrates mastery. If a student answers every question correctly, he completes instruction in

much the same way as if he were in a linear program. However, each time he makes an error, he must complete a remedial sequence before being returned to the main stream of instruction. Further, different errors necessitate different sorts of remediation, and these in turn may lead to different amounts of drill. The major distinction, then, between the linear and the branching approach is that while student error is discouraged in the first instance, it is used as a diagnostic tool in the second. Such a diagnostic procedure makes it possible--at least theoretically--to tailor instruction according to the particular needs, abilities, and learning style of each student.

When the branching strategy combines with the computer, the result is a dynamic, interactive CAI program. It is the computer's great capacity to store and retrieve the information required to individualize instruction far more efficiently than printed branching texts or "scrambled texts" that makes it a potentially powerful instructional tool. Moreover, a number of researchers have found the motivational effects of CAI are very great for some learners, and much of the time these are remedial students--the population to which the materials of this study were directed.

One can view the printed linear text and the CAI program with D&P as being at opposite ends of an instructional continuum. The linear mode gives only the basics of instruction which CAI offers extra remediation, specifically designed to respond to differences in learners. In between these extremes, there are at least three other possibilities: 1) a printed

branching text without D&P; 2) a printed branching text with D&P; and 3) CAI without D&P. Numbers 1 and 3 are identical in content, as are number 2 and CAI with D&P. Their respective differences are the way in which instruction reaches the student--via book or a computer terminal. (Although it would have been possible to produce a computer program with the same content as a printed linear text, we felt that the linear approach on a computer was too trivial to be of interest.) Two essential differences exist among the five modes used in this research: the amount of instruction available and the way that instruction is delivered.

One would expect that--other things being equal--the more instruction a student gets, the greater his achievement. Further, one would also expect that a dynamic and interactive delivery system would motivate students even more than PI textbooks. In theory, at any rate, CAI with D&P should outperform all other modes because it provides ample instruction tailored to individual needs and presented in a highly interactive fashion. The student using a linear textbook theoretically will suffer from the absence of both fully-individualized remediation and instruction through a more intrinsically dynamic medium. And students in the other three modes should fall somewhere in between these two points in terms of their achievement.

PROBLEM STATEMENT

Whether such computer capacities are really significant factors in learning leads to the formulation of the following problem statement:

> Does the use of CAI (with its dynamic branching sequences) provide more effective and/or more efficient learning of the mechanics of writing for a given population than a printed linear or printed branching programs presenting the same material?

The significance of this problem lies in the great difficulties which a mass educational system has in providing high standards of instruction to large numbers of students. This problem appears especially acute in the area of remedial language skills, the subject matter of this experiment. The decline in literacy in the past ten years has been precipitous and commented upon many times. As we have said, financial realities make it impossible to offer the sort of individualized attention the problem demands, and even if this were possible, disagreement exists among educators about the best way to teach writing. Most frequently, there seems to be a confusion between teaching "style" and "mechanics"; this research began with a crucial distinction between these two skills. Style--a vague term suggesting the most appropriate expression of ideas--is a function of intellectual maturation. It involves so many complex cognitive, affective, and psychomotor operations (which legitimately combine and permute as endlessly as there are writers with "good style") that an adequate operational definition cannot be produced readily. Further, there exists no coherent and empirically valid learning theory that accounts for the way in which individuals acquire the skill of "good style". Accomplished stylists can be produced only after a long period of time during which they read widely, practice writing continually, and receive close and personal supervision.

In short, the teaching of good style appears beyond the capacities of mass education. Certainly, the usual one-semester course in freshman composition is woefully inadequate.

The teaching of mechanics, however, differs from the teaching of style because one can measure grammatical correctness and the like against a recognized criterion level--the rules and conventions of Standard English. (There is no shortage of Freshman composition texts, each with its own stylistic advice, but they are remarkably consistent when it comes to what constitutes "standard English" i.e., grammar, syntax, and mechanics. See Richar Corbin et al, Guide to Modern English or Langdon Elsebree et al, Heath's College Handbook of Composition as typical examples.) Such rules can be made as "true" (or at least as rigorous) as those of mathematics, and so taught. This does not mean that "creativity" (again, a word that resists operational definitions) should be abandoned in favor of some mechanistic rote learning of rules to be followed slavishly. But "creative writing"--even basic expository prose--is impossible without adequate mastery of fundamental language skills. This research assumes that there are various principles in the mechanics of writing that can be enunciated and applied to various utterances. Once these principles have been specified, they can be incorporated into PI or CAI programs. The results of this are potentially powerful. Teachers can be freed of the tedious and often wasteful task of day in and day out teaching and drilling of basics. They would, in theory at any rate, be able to concentrate on the more complex skills

involved with style which cannot be so readily programmed. A significant saving in money and time should be effected, but the courseware that will teach mechanics of writing still resides very much in the developmental stage. It was the further development and evaluation of such material--and the methods of delivering it--on which this research focused.

THE STATE OF THE QUESTION

Although theoretically CAI with its branching sequences should provide more effective learning, most studies comparing CAI and linear texts have not been conclusive. (Studies comparing CAI with "traditional" teaching are *ipso facto* invalid, since the variables of the "traditional" classroom cannot be controlled.) "No significant difference" appears to be the most frequent conclusion in comparisons between PI and CAI programs. Where differences exist, they are often those of motivation or a more rapid achievement of objectives in favor of the computer. Caldwell (1973), Brown (1969), Mull (1973), Hong (1972), and Johnson (1973) all report students preferred the computer to the printed text. It should be noted that Lenn and Maser (1971) claimed CRT (cathode ray terminal) used with a CAI used for more than half an hour actually tired their subjects, the result being that the printed treatment was more favourably received. However, the motivational potential of the computer must still be seen *as* potential because of its novelty in the classroom. Perhaps only when the computer becomes as familiar as textbooks (and sufficient data generated)

will we determine whether or not the motivational effects of computers are as genuine as early studies claim. Sufficient data to overcome this novelty effect has yet to be generated; this experiment recorded data on student attitudes in order to add to the available information.

The saving in time frequently reported by CAI programs has been, in some cases, rather misleading. Studies comparing CAI with traditional teaching claimed savings of up to fifty percent (e.g., Homeyer, 1970); but one well-controlled experiment, Dick and Lippert (1968), showed a savings of between only 12% and 17%--significant, but not as dramatic. In studies comparing PI and CAI, however, the results are far more mixed. Grubb and Selfridge (1964) and Aboud and Bunderson (1973) record findings favorable to the computer. But Dick and Latta (1971), Dick et al (1974), and Caldwell (1973) find in favor of PI.

Investigators often are most concerned with differences in overall achievement. Here again, results are inconclusive. McMurray (1964), Lenn and Maser (1971), Bossone and Weiner (1973), Roid (1971), Dick et al (1974), Caldwell (1973), Brown (1969), and Blitz (1972) all reported no significant differences in achievement. Only two studies, Grubb and Selfridge (1964) and Aboud and Bunderson (1973) show results in favor of the computer, while Dick and Latta (1970), Mull (1973), and Hong (1972) found superior results with linear texts.

It may be of some value to look more intensively at a few typical studies. We have chosen three such studies, all of which compare achievement

using CAI and PI, but with quite different results. The first finds in favor of PI, and the third finds no significant differences.

The first study we would like to look at is "Effects of Branching in a Computer Controlled Auto-Instructional Device" (Coulson, et al, 1961). Although somewhat dated and using terminology out of current fashion, the experiment appears well controlled and clear in its objectives. The researchers intended to explore the effectiveness of a branching program on a computer (which they called "variable-sequence instruction"); the essential feature of this instruction was that branching decisions were made on the basis of student error and the student's own evaluation of how well or poorly he was doing. The hypothesis that the study tested was that students in such a program would score higher on a posttest than another group in a printed, linear program (which they call "fixed-sequence instruction"). Further, it was thought that the CAI group would complete instruction in less time.

Subjects were thirty high school students, divided into two matched groups. (The authors do not comment upon the fact that all were paid volunteers nor on the effect such a sample might have on the study.) The course material attempted to teach principles of logic, and 345 items were prepared, of which 233 were chosen by an outside consultant for use by the PI students. The consultant organized these items into a sequence which "he felt [our italics] would provide effective instruction". The branching students, of course, had the entire group of 345 items as frames

to which they could be branched, and branching proceeded on three bases: 1) the cumulative number of errors on each topic; 2) the diagnosis of misunderstanding as indicated by responses to individual questions; and 3) learner control (whether to review or to go on to new material). The content was so written that the first contact students had with a concept (referred to as "the main stream") was presented in fairly large increments, while the remediation items on the same concept were smaller and more numerous.

The posttest consisted of 95 questions, about half of which were multiple choice, and the remainder "fill in the blanks". The results showed that the CAI group did significantly better: at the .05 level ($t=2.15$, $df=28$), their mean score on the posttest was 74.3, compared with 65.4 for the PI group. The mean time in minutes required to complete instruction was 116.2 for the CAI students and 129.5 for the PI group. This difference in training time was not found significant.

The authors conclude that the new CAI program gave a more accurate diagnosis of student needs and more effective remediation than either the old CAI or the PI programs used in the second study. Low ability students performed as well as high ability ones on CAI, although they did need more time to reach that level. However, low ability students on PI also did as well as high ability PI students; thus branching alone cannot account for the good scores of the lower ability CAI students. The significant differences in posttest scores result directly, conclude the

authors, from CAI's sharper diagnostic procedures and from the effectiveness of the remedial material.

A second study found in favor of PI. "Comparative Effects of Ability and Presentation Mode in Computer-Assisted-Instruction and Programmed Instruction" (Dick and Latta, 1970) followed sixty-four low and high ability high school students through a short course in mathematics. Half the subjects (represented equally by low and high ability students) worked through a printed, 52-frame linear PI text, and the other half used an "enhanced" version of the same material. By "enhancement", the authors mean special remediation hints and the chance to look back and review. Only four additional frames, not available to the PI group, were written for the computer. A 39-item posttest (which had a .79 reliability on the Kuder-Richardson scale) was used for both pre- and post-testing. Before the experiment began, teachers gave the low-ability subjects special help in the terminology they would encounter during instruction, and all subjects worked through an 8-frame PI lesson on another mathematical concept in order to familiarize themselves with the basic format. CAI groups were given a brief explanation on the use of the terminals, and then bussed to the computer center on the day of the experiment. The PI groups stayed in their own classrooms. (Effects of travelling to a new environment were not considered by the authors.) Upon completion of the instruction, all students were given posttests.

The researchers wished to test the hypothesis that low ability students

would do better on CAI than PI because of CAI's diagnostic and remediation capacities. Further, by using CRT's, they hoped to learn about the dynamics involved in learning from a terminal that did not provide hardcopy.

The data collected consisted of the scores on the posttest, scores on a retention test given two weeks later, the number of errors made, and learning time. The results showed that the PI group did better overall, although there was no significant difference between the two high ability groups. Low ability CAI students, however, did significantly worse than any other group, including low ability PI students. An analysis of various items on the program did not reveal any special problems that would account for this, nor did the variation between constructed or recognition responses appear to contribute to the poor showing by low ability PI students. Low ability PI students not only scored better than their CAI counterparts, but also showed a significant saving in learning time. Thus, the group most often assumed to benefit from CAI--low ability students--fared far worse than what had been expected. The authors assume that the continuous flow of information on the CRT was too much to be assimilated by low ability students. They suggest that the control the learner has over review with a printed text should become a feature of CAI instruction, especially when using a CRT. (One might assume that hardcopy would overcome this problem of review, but apparently such is not the case: another study showed that CAI is not necessarily "enhanced by the provision of a feature for the production of hardcopy" (Dearborn

and Calingaert, 1972).

A third study found no significant differences between CAI and PI. "Computer Assisted/Programmed Instruction on the Law" (Lenn and Maser, 1971) used materials developed for police officers, and drew their subjects from both recruits and men taking inservice training. The authors found both CAI and PI "effective" and well received by subjects. Students worked on the PI material for fifteen minutes a day for four weeks, and the difference between pre- and post-test scores was about 19%. (One weakness of the experiment is its failure to be very specific about its data--hence "about 19%".) The same material taught via the computer was "found to be equally effective as the printed material" (p.2). However, CAI took longer for the students to complete. (Again, actual time is not reported.) This may well have been because some of the CAI remediation referred students to an adjunct printed text. The computer did appear to have a moderate motivational value which the authors attribute to "novelty". CAI was also judged easier than PI to update instruction and to maintain records. But costs (which have since fallen dramatically) were too high the authors felt to warrant implementation, especially since final achievement was about the same for both groups.

From the authors' description, their PI and CAI materials appear to be quite like most other programs. The PI version was linear, permitting the student to vary only the pace of his progress. The CAI program used student responses to decide whether to review or go on, to provide a second chance on a missed question, or to receive a remedial comment.

However, this particular computer's inability to accept anything but exact matches led the authors to develop a self-testing procedure whereby students wrote their answers on pieces of paper, and decided themselves how close they were to the answers the computer would display. Students then typed either "yes" or "no" ("yes" indicating their answers were correct) and the computer would branch accordingly. The authors judged that such branching was "time consuming and not generally worthwhile": PI students were able to complete material in 2.5 - 4.0 minutes but CAI students averaged 5.0 - 6.0 minutes. This delay, the authors write, occurred because of the computer's delay in responding with remediation and writing out new instructions; further, there was a "possible" distraction caused by the poor quality of print of the CRT used. Typing the answers also was found to slow down the whole process. That all these problems are merely technical ones which can be relatively easily eliminated was not considered in the study.

The study began with the hypothesis that the "novelty and liveliness" of CAI would increase motivation. In fact, the terminal "tired" students, many of whom requested rest periods after about half an hour. Students also preferred the selfscoring mode used by PI. Finally, students favoured PI because it provided a better chance to review, thus resulting in better "self pacing" and "self managing". (Again, these are technical problems relatively easy to solve.) The authors do concede, however, that their sample was too small to make any definitive judgements.

We are not suggesting that the three experiments we outline here are anything more than a brief sampling of the literature; nonetheless, we think them representative. They do point to some of the most frequently-cited issues in the literature: does CAI provide a sharper diagnostic instrument because of its branching capacities and do these branching capacities provide more effective remediation; does the flow of information in a CAI lesson come too quickly to be assimilated and therefore does a printed text offer better student control over the rate of instruction; does CAI increase motivation or does the self-scoring feature of PI make it the favoured mode; does CAI result in significant time saving? However, we think that none of these studies recognizes adequately variables other than the actual delivery system. There are surely differences beyond the fact that one mode branches where another doesn't or that either the student or the machine turns pages. Such oversimplification of research questions may explain, in part at least, why the literature is so inconclusive. The right questions are not always being asked, and often a reader would like to know more about what factors other than the delivery system were at work. To this end, this experiment considered the presence or absence of branching as an added factor, and our research design reflected this.

One is faced, therefore, with puzzling literature. The theoretical advantages of CAI have yet to be demonstrated in practice. Rockert and Norton (1975) put the situation neatly:

> In summary, the results in the drill-and-practice and tutorial modes of instruction are ambiguous at the university level. It does appear that there is evidence to support their use for structured, repetitive kinds of material, although there has yet to be a demonstrable cost-effectiveness finding that would support their use in replacing instructors at the college level. Possible misapplication of the technique vis-à-vis substantive course material, plus the general failure to do research on the impact of such systems, has resulted in a feeling of disenchantment. Little new work except in remedial learning is reported with regard to initial presentation of the material, and no supportive evidence is available in the literature. (p.104)

The content of this experiment--remedial English--falls within the area that Rockert and Morton think most promising. The mechanics of writing can be structured and require much repetition for mastery. However, the literature does not show sufficient examples of controlled studies comparing remedial English taught by CAI and a printed linear text.

OBJECTIVES

The positive hypothesis for this experiment can be stated as follows:

> A CAI program in the teaching of the mechanics of writing to a given population results in more effective learning (as measured on a standardized test) than a parallel, printed linear version.

As the problem statement indicates, CAI (as the term is used here) involves branching the student to different portions of the program, contingent upon responses made. Hence the sequence and rate of learning varies with the individual while a linear program varies only the rate.

The primary objective of this research was to demonstrate that the branching capacities of CAI are responsible for its superiority. Two subobjectives also exist. One was to explore the relationship between the amounts of D&P and overall achievement. The second was to demonstrate that a computer is a more efficient means of presenting branching sequences than a scrambled text, the printed equivalent to CAI.

DEFINITIONS

Linear Program: A printed text in which material is broken down into its constituent parts, placed in frames containing relatively small units of instruction, and which require the student to make a response to each frame. The student receives feedback in subsequent frames, but can use such feedback only to vary the rate at which he completes the instruction.

CAI: A computer program so constructed as to require no special computing skills by the student, and which presents instruction in a tutorial mode, tests to see if mastery has been achieved, and directs the student to the appropriate remediation or D&P sequence. The student can vary the pace, the sequence of instruction, and the amount of D&P.

Branching Sequences: The sequence of frames to which a student is directed on a CAI program appropriate to his needs, and which are contingent upon the responses he makes. A branching sequence may contain remediation, D&P, testing, or presentation of new material.

Drill-and-Practice: A number of questions which provide the student with additional practice in a specified concept or skill.

Effective Learning: The ability of the student to apply the principles enunciated in instruction to sample sentences, as measured by scores on a standardized test.

The hypothesis may be restated as follows:

> Students who go through a CAI sequence with branching sequences will perform significantly better on a standardized test than students learning the same material in a printed PI text, whether linear or scrambled. Students who go through a CAI lesson with drill and practice sequences as well as branching sequences will perform better on a standardized test than students learning the same material through a CAI lesson with branching sequences only.

POPULATION

The subjects of this experiment were 75 remedial English students

enrolled in first-year courses at Vanier College, Snowdon Campus. A remedial student was defined as one who had scored in the lower fifth of those students taking the Missouri College English Test (Callis and Johnson, 1964). This standardized test was given to all first-year students during September 1977; some students were placed in remedial writing courses as a result of this test, and they were excluded from the study. Eighty-two of the remaining students were enrolled in English-190 and English 404, remedial writing courses which began in January 1978. There were four sections of these courses, with approximately twenty students in each one. These students acted as the five treatment groups.

PROCEDURE

The experiment randomly divided the original 82 students into five treatment groups: 1) printed linear text; 2) printed scrambled text; 3) printed scrambled text with D&P; 4) CAI without D&P; and 5) CAI with D&P.* (There was a mortality rate of 7 students, reducing the final total of the population to 75.) Material was presented to each group only through the designated treatment. The classes met for the standard three hours a week with an additional one hour being set aside for work on the specialized material. In the case of the students assigned to the two CAI modes, this time was spent as a regularly scheduled hour at the terminal. The other three groups were permitted to select their hours to work on the printed texts. Classroom time was devoted largely to developing skills other than those taught by the five treatments; thus a significant amount of instruction was given on logic, organization, argumentation,

*We identified the groups as follows: 1) L; 2) B; 3) XB; 4) C; and 5) XC.

and the correction and review of student work.

Each week, students completed one or two lessons and wrote the criterion test at the completion of each module. Because the printed lessons were distributed in class without the kind of monitoring afforded by the CAI scheduling, the teachers were unable to know if the PI students were or were not working their way steadily through the lessons and writing the criterion tests, although so instructed. It was possible to collect data only on the performances of the computer groups as they worked their way through the lessons; and therefore, this data was not used in making the final evaluation between modes. Further, we assumed that PI is by nature more selfpaced than CAI where the students have only so much time to use the terminal. Monitoring PI groups would have, we think, somewhat distorted the PI experience because this prccedure would not necessarily go on in an ordinary course using PI materials.

Monitors were present at every CAI session to ensure that students had no difficulty operating the equipment. They did not, however, answer any questions pertaining to lesson content. The presence of monitors was necessary because the system did not have the features that would have permitted students to work completely independently (for example, their own accounts, a start-restart function, automatic recording of scores and relevant data into a separate and cumulative grade file). The presence of the monitors was necessary, therefore, not because of intrinsic drawbacks of CAI but because of deficiencies in our system. We hope to remedy this in the future.

As we have said, all groups worked independently, the major difference being that computer groups were scheduled at particular hours each week. This was done to ensure that the limited terminal facilities be made available to students on a fair and equitable basis. For the first eight weeks of the experiment, this presented no difficulties. However, labour problems at Vanier, which closed the school for five weeks, necessitated an accelerated schedule when classes resumed. It is difficult to assess the effect of such acceleration. At any rate, the paper groups also had to work their way through the remaining texts at a faster pace.

The computer groups encountered a number of problems similar to those cited in many other studies. The most serious was the breakdown of the timesharing system; such system "crashes" occurred about two or three times a week, and varied from a few minutes to about an hour. This meant that students had to come back for makeup sessions. Much worse, however, were those times when the system failed just as a student was in the middle of a lesson. The data accumulated on that student's progress would be lost, and there were not facilities either to restore this data or to permit the student to restart at the point of the system breakdown. This may account for the relatively large amount of frustration expressed by the CAI groups in their overall assessment of the program (see Table VI--Attitude Survey). Such problems are, of course, all too common in CAI in which timesharing is used. Moving to minicomputers would be one solution; but, given the present state of the hardware, it would require giving up a considerable

portion of the computing power available on large timesharing machines.

The printed texts had deficiencies of their own. The PI books were reproduced from the computer listing and so were distributed about a week after the computer lessons were ready to be used. Restrictions on time and funds meant that production values for the books were less than we would have wished. Simply because of the problems of producing and correcting masters (with one master for every page), there were more typographical errors in the printed versions. And physically, the books--which were mimeographed and stapled together--lacked visual clarity and durability. However, as the attitude survey indicates, students did not seem to consider these problems as major issues.

Most students in the computer groups took about thirty minutes to an hour to complete a lesson. As one might expect, the computer D&P group finished somewhat more slowly than the group without D&P. But students did not complain about feeling pressures of time. It was not possible to gather data for the paper groups on time spent since they generally worked on their own and often at home. However, they too did not vocalize complaints about time spent (see Question 22 of the attitude survey). In fact, it is our feeling that one of the usual measures in this kind of research--time to complete the work--is not an important issue. We were far more interested in students achieving mastery than in the amount of time they took.

At the completion of the lessons, all students sat for a posttest based

on three standardized tests (The McGraw Hill Writing Skills Test, The College English Placement Test, and The Missouri English Test). We will review the problems associated with choosing an appropriate test instrument in the Discussion section of this report. In brief, complete sections of these three tests were removed and reproduced in a single format. Care was taken to adhere to the time restrictions for the original tests. Some editing was necessary because no commercially available standardized test adequately covered the material in the lessons. Thus one test might assess a student's logical skills, another spelling, and a third his stylistic problems--none of which were part of the instruction given the treatment groups. In addition, we collected data in the form of essays students submitted during the course of the semester. This was further broken down into an overall score and the average of the last three assignments. We report on the results below.

THE MATERIAL

The material for all modes was written during the first semester, and revised, produced, and distributed during the course of the second semester. In all, twelve lessons were written. They are:

Lesson Name	Content Covered
CINTRO	The comma to separate introductory elements from the rest of the sentence;
CPAIR	The comma to enclose nonessential elements of a sentence;

CSER	The comma to separate items in a series;
CMAIN	The comma to separate two main clauses joined by a coordinating conjunction;
NOCOM	The conditions governing the omission of the comma;
APOST	The use of the apostrophe;
DANGLE	Identification and correction of dangling and misplaced modifying elements of a sentence;
VERB	Verb tense and agreement;
PRCASE	Pronoun case;
PROGRE	Pronoun agreement;
CAPI	The principles of capitalization;
REVIEW	A fifty-question summary of all other lessons.

Full documentation of these lessons is included in the "Remarks" lines in the computer listings which accompany this report. The following is to inform the reader about the general strategies employed in the lesson design, rather than provide an exhaustive discussion of particular lessons.

It became clear as the lessons were being written that no absolute hierarchy or entailment sequence of skills could be followed. The skills taught by the lessons are essentially coordinate; each lesson is quite self-contained, and requires no prerequisite knowledge that is taught by any other lesson. For instance, a student would not have to have gone through CAPI to succeed on PROGRE, nor would his mastery of PRCASE be contingent on his mastery of DANGLE. The one exception to this might be REVIEW which does

indeed review all other lessons. However, we believe that REVIEW can readily be renamed PREVIEW and used as a diagnostic instrument. Our experimental design did not allow a pretest, but we intend to produce just such a lesson in the future.

The absence of a hierarchy among the lessons does not imply that there are no prerequisite skills needed to do them. Rather, we assumed that the prerequisite skills the lessons demanded were those possessed by entering collegial students--even remedial ones. Typically, these skills included being able to recognize the sentences, knowing the meaning of words like "verb" and "noun", and other such basic concepts. However, a fundamental strategy of the lessons was that each time a term--however basic--was introduced, we would include a very brief review. (For example, VERB says to the student: "As you know, a verb is a word that describes an action.") But when the terminology is likely to be less familiar, the courseware goes into somewhat fuller detail. (CMAIN, for example, discusses the coordinating conjunction.) We tried to find a middle path between writing large amounts of exposition (which tends to diminish the interactive nature of the lessons) and ensuring that no student would have trouble with the lessons because of terminology. As the attitude survey indicates, 87% of the population felt that the pace of the lessons was about right, and 89% said that the degree of difficulty in the lessons was also about right. We take this as confirming our estimates about the amount of prerequisite skills the lessons demanded as well as the appropriateness of the lesson

content generally. We think, therefore, that the lack of a hierarchy, insofar as it permits a flexible sequence for the student to go through, is desirable.

Each lesson follows the same basic strategy of presentation. A brief opening statement provides the student with a set of behavioral objectives. Although each lesson contains one major concept, it is broken down into subconcepts. Like the main concepts, these subconcepts teach coordinate skills, and therefore, there is no absolute requirement that they follow a particular order. Each subconcept is introduced by means of a somewhat abstract general statement, quickly followed by a sentence which applies the principle under discussion. The student is then given a sentence of his own in which he has to apply the principle. This first question is written so that in case of an error, the fullest amount of remediation is immediately offered. (This is not true in the linear mode which has no remediation.) Further questions are generated, each one containing remediation frames for wrong answers. In fact, at the outset of the presentation of a concept, even correct answers are given a brief statement of why the answer is correct. This mitigates the effects of students simply guessing. In CAI with D&P, a single wrong answer in a particular subconcept automatically results in the student going through an additional series of questions on the subconcept. These D&P sequences contain about three or four questions, each one having its own set of remediation frames as well. After all the subconcepts have

been covered, the lesson ends with a criterion test. Here, the student is told only if he is right or wrong and his score is tallied.

As we have said, each lesson follows the strategy outlined above. However, the nature of the different modes means that there are differences in presentation. Obviously, a student in the paper versions must turn to the proper pages, observe an honour system in terms of looking ahead at right answers, and tally the score himself. The students in the scrambled text with D&P mode are only told to do the extra questions, not automatically placed there. Moreover, the presence or absence of D&P--whether PI or CAI--obviously means different amounts of instruction. This is most striking when one simply compares the sheer bulk of the linear text (which contains only the basic instruction and questions) with the scrambled text with D&P which contains basic instruction, questions, remediation, as well as D&P. While the essentials of instruction were identical, the nature of the delivery system necessitated changes in how the questions were presented. Once again, we direct the reader to the courseware itself for concrete examples.

The following chart shows the basic strategy for the lessons:

1. STATEMENT OF OBJECTIVES PRESENTED TO THE STUDENT
2. INSTRUCTION IN FIRST SUBCONCEPT
3. SAMPLE APPLICATION OF FIRST SUBCONCEPT
4. TEST OF FIRST SUBCONCEPT (SEVERAL QUESTIONS)
5. MASTERY OF QUESTIONS? YES--GO TO NEXT QUESTION
   NO--REMEDIATION

6. MASTERY OF SUBCONCEPT? YES--GO TO NEXT SUBCONCEPT
   NO--D&P
7. REPEAT PREVIOUS STEPS WITH NEW SUBCONCEPT
8. CRITERION TEST
9. RECORD RESULTS

The reader, of course, will recognize that D&P is only possible with two of the modes, and that no remediation is available in the linear mode.

There are also other differences between modes, perhaps not so readily apparent. The printed lessons were reproduced from the computer printouts; and despite careful checking, a number of typographical errors were not caught. Moreover, it was not always possible to replicate the style of the original computer lesson. For instance, if a student in CAI was asked to identify an error in a particular sentence, he would simply type in his answer, and the computer would scan the answer and reply accordingly. In the scrambled texts, which tell the student to turn to a particular page for each response, we were faced with the problem of inadvertently prompting the student by selecting certain words in the sentence for him to turn to. Otherwise, we would have had to have a page listed for every single word in the sentence, a clearly unworkable feature. We were not able to assess how much of an effect this had on learning. Nonetheless, we think it important to point out that although material was identical about 95% of the time, some modifications were necessitated by working in a printed as opposed

to a computer-assisted medium.

One of the most important features of the CAI--and one which was impossible to replicate in any of the printed versions--was the automatic scorekeeping function. The system recorded and stored every answer the student made as either 1 or 0 (for right or wrong). This meant that a tally was ready immediately upon completion of the lesson. But even more importantly, the student's scores were broken down into subscores that reflected mastery on subconcepts, D&P sequences, and criterion tests. A printout was generated at the completion of each computer lesson which showed precisely which items were answered correctly or incorrectly, what degree of mastery was reached for every concept, whether or not the D&P sequences were effective, whether the criterion test scores were commensurate with the student's work elsewhere in the lesson, and of course, the final score on criterion test. Since our experiment was limited to comparing the effectiveness of various modes, we have not analyzed this data. Next year's work, however, will begin with a detailed analysis of the results of all CAI student scores. This will be central, we think, to our future revisions; defective questions will be eliminated, effective ones used as models for new material. One especially interesting possibility would be to generate data reflecting how clusters of questions have been answered. For example, we could examine how constructed response questions compared with multiple choice questions or the degree to which the population correctly answered the first, second, or subsequent questions on a particular

subconcept. High on our priorities will be the creation of a file system so that such data can be deposited directly into a separate scorekeeping file for later analysis.

We also believe it is important to note that labour problems occurred at Vanier during the sixth week of our research. Approximately 80% of the students had completed all CAI lessons when the College was closed by a strike which lasted for five weeks. Concordia University offered both their computer and classroom facilities so that the students could complete the program without interruption. Unfortunately, it proved impossible to arrange for more than about 60% of our treatment populations to participate. Rather than compromise the data, which would have been based on very different conditions, we decided to wait until the strike was over to continue so that all subjects would be tested under the same circumstances. Although we cannot fully assess the effects of the strike in terms of final test scores, we can say that our research design ensured that these effects were felt equally throughout the test population. Because our primary objective was to measure the relative value of five different modes of delivering instruction (rather than measuring pre- and post-treatment gains), we think our results are valid. However, we also recognize that the posttest after a five-week delay may not yield fully generalizable results. Although we have only anecdotal and impressionistic evidence, we think that our students did regress somewhat during the strike, and that an uninterrupted experience would have been decidedly preferable.

We can only repeat that because subjects were randomly assigned to treatment groups, the negative effects should be felt equally throughout. One other point, however, should be made. The mortality rate for the experiment was felt only after the strike; all of our dropouts (about 7 of 82) participated up to the strike.

Finally, we should report that we experienced no serious problems in terms of student ability to handle the terminals. Very little time--literally a few minutes--was required to instruct students. Only about 5 to 10% of regular sessions on the terminals were missed (often by the same three or four students), and most were made up promptly. In short, except for the system failures, there were few if any serious logistical problems encountered.

DESIGN

The experiment can be diagrammed as followed:

R1	X1	O1
R2	X2	O2
R3	Y1	O3
R4	Y2	O4
R5	Z	O5

where X1 is the computer mode with D&P, X2 is computer without D&P, Y1 is the scrambled text with D&P, Y2 is the scrambled text without D&P, and Z1 is the printed linear text.

DATA AND INSTRUMENTATION

Data was based on the following observations:

1) A 70-item posttest combining complete sections of The McGraw-Hill Writing Skills Test, The College English Placement Test, and The Missouri English Test. These three standardized tests were edited to reflect the materials actually taught by the lessons, with inappropriate questions removed and time constraints readjusted accordingly; 2) The average of all essays submitted during the course of the semester; 3) The average grades of the last three essays submitted during the semester; 4) A 26-item attitude survey.

RESULTS

The relative achievement of the five treatment groups was measured by means of a one-way analysis of variance of test scores, average grades for all essays submitted during the semester, and the average grade for the last three essays submitted. These results are shown in Tables I. II. and III respectively.*

TABLE I ANALYSIS OF VARIANCE OF TEST SCORES FOR THE FIVE TREATMENT GROUPS

VARIATION	DF	SS	MS	F TEST
TOTAL	74	7805.28		
TREATMENTS	4	389.968	97.4919	.920317
ERROR	70	7415.31	105.933	

* One student from group XB and one from group B completed the lessons and sat for the posttest, but withdrew before completing their essay assignments. Hence the discrepancy between the number of subjects in Table I and in Tables II and III.

TABLE II ANALYSIS OF VARIANCE TABLE COMPARING SCORES ON ESSAY AVERAGES FOR THE WHOLE SEMESTER

VARIATION	DF	SS	MS	F TEST
TOTAL	72	2156.16		
TREATMENTS	4	26.0662	6.51656	.208031
ERROR	68	2130.1	31.325	

TABLE III ANALYSIS OF VARIANCE TABLE FOR AVERAGE SCORES OF THE LAST THREE ESSAYS

VARIATION	DF	SS	MS	F TEST
TOTAL	72	2383.37		
TREATMENTS	4	84.9074	21.2268	.627996
ERROR	68	2298.46	33.8009	

The F-scores for all three tests were not significant. Results are similar for all measures. In order to ascertain whether the results obtained here were from populations with normal distributions, the means and standard deviations for all groups were examined. The results are reported in Table IV.

TABLE IV THE MEANS AND STANDARD DEVIATIONS OF TEST SCORES FOR THE FIVE TREATMENT GROUPS

GROUP	N	MEAN	S.D.
Linear	17	65.13	10.17
Scrambled without D&P	14	58.37	7.03
Scrambled with D&P	18	63.33	9.42
CAI without D&P	12	60.86	19.37
CAI with D&P	14	61.23	11.97
All groups	75	61.35	12.38

Results indicate a normal distribution.

Because the sample population included a number of students whose previous education was not entirely in English, a Chi-square analysis was performed to test whether any treatment group had a disproportionate number of such students. The results are reported in Table V.

TABLE V CHI-SQUARE ANALYSIS OF DISTRIBUTION OF SUBJECTS WITHOUT ALL PREVIOUS EDUCATION IN ENGLISH

	B	XB	L	C	XC
Non-English	4* (3.1)**	5 (4.0)	4 (3.8)	2 (2.9)	2 (3.1)
English	10 (10.9)	13 (14.0)	13 (13.2)	11 (10.1)	12 (10.9)
Total	14	18	17	13	14

Chi-square.....1.519
Degrees of Freedom.....4
(Not significant)

No significant differences were found between groups.

An inventory of student attitudes about the different modes was conducted. The results of 18 key questions (that is, those that dealt directly with the delivery systems rather than content) are shown in Table VI. The survey itself is reproduced in Appendix I.

* Observed frequencies
** Expected frequencies

TABLE VI PERCENTAGE OF FAVOURABLE RESPONSES ON ATTITUDE SURVEY

QUESTION	ALL GROUPS	XC	C	XB	B	L
1*	NA					
2*	NA					
3*	NA					
4	64	64	77	61	65	59
5	72	57	60	89	76	71
6	46	43	46	45	50	47
7	75	78	54	88	64	88
8	68	64	38	78	71	82
9	77	71	61	89	71	88
10	70	64	54	67	86	76
11	70	88	70	77	64	53
12	62	64	54	78	50	59
13	76	79	77	71	79	76
14	71	86	70	61	71	71
15	82	93	77	72	71	94
16*	95	100	93	100	93	88
17*	93	86	100	100	93	88
18*	41	43	46	22	57	47
19	72	93	85	67	57	65
20	71	93	69	78	64	53
21	82	78	85	83	78	82
22	70	64	84	78	50	83
23	58	50	61	61	43	71
24	67	57	77	72	50	77
25*	87	92	85	89	86	82
26*	89	86	100	76	93	94

*These questions pertain directly to content or language of previous education.

In Table VII, we report on the one-way analysis of variance performed on this data. No significant differences were found.

TABLE VII ANALYSIS OF VARIANCE FOR PERCENTAGE OF FAVOURABLE RESPONSES ON ATTITUDE SURVEY

VARIATION	DF	SS	MS	F TEST
TOTAL	89	16462.9		
TREATMENTS	4	878.511	219.628	1.19789
ERROR	85	15558.4	183.346	

The results suggest that there are no significant differences in achievement or attitude among all groups. The random assignment of population to various treatments showed that no mitigating factors such as previous education in English made any difference in the data collected. The means and standard deviations for all groups were very similar. Hence, we cannot attribute the lack of significant results to any anomolies in the sample. The test scores and the two sets of essay scores also show no significant differences. If the posttest had lacked content validity, we might have found marked differences in results between test scores and essays. This was not the case. There is only slight evidence that the CAI group with D&P was doing better as the semester progressed; the mean scores for the final three essays of the DAI with D&P did improve. In fact, their mean score went from being the lowest ranked on the test results to the highest on the last three essays. However, these results are not any more statistically significant than any other measure in the experiment.

The results of the attitude survey show that all groups expressed favourable responses to the program, regardless of delivery mode. Scores on the survey were regrouped into "favourable" and "unfavourable" for 18 key questions. In looking at these statistics, there were some differences of note. Eighty-six percent of the CAI group with D&P considered theirs an "enjoyable" experience as compared to 71% for all groups (and a low of 61% for the scrambled texts with D&P). Ninety-three percent of them also thought that they learned more efficiently compared to a group average

of 72%. However, 43% of the CAI with D&P group (and 40% of the other CAI group) reported frustration at times, compared to a group average of 28%. This can likely be attributed to the system malfunctions. All groups reported very favourably about the appropriateness of the pace (87%), level of difficulty (89%), and clarity (95%) of the lessons. This indicates a high acceptance for the materials themselves, regardless of the way they were delivered.

DISCUSSION

Our positive hypothesis was that a CAI Program in the teaching of the mechanics of writing to a given population results in more effective learning (as measured on a standardized test) than a parallel printed linear or scrambled version. The results of the experiment show no significant difference in achievement among the treatment groups. This holds true whether one examines test scores, average grades for all submitted essays, or averages for the last three essays only. Scores on all three measures were remarkably consistent; only CAI with D&P markedly changed rank among the treatments for the final three essays, but the mean scores here were not statistically significant.

We do not take the consistency of these results as definitive because we are not convinced that our posttest was sufficiently comprehensive and discriminating. Aiming for maximum objectivity, as well as recognizing our own limitations in devising a fully reliable measuring device, we chose to use a hybrid of three commercially-available standardized tests.

But this new test included only four categories of errors (punctuation, grammar, capitalization, verb) and a "no error" response--a constraint imposed by the opscan sheet answer format. The lessons themselves had made many more discriminations. Secondly, the test required only a recognition of an error via a multiple choice question, not a recognition and a correction. Again, this was different from the lessons themselves. In fairness, the test did have the virtue of being standardized, readily administered, and having a format that would favour no one treatment group. Moreover, the test questions themselves had been the basis for much of the content of the lessons, even though the styles of presentation were dissimilar. Thus, we feel that, given the lack of the kind of commercial standardized test we wanted and the constraints of time that made it impossible to construct our own reliable test, we made a realistic compromise.

However, we also felt it necessary to look at essays as well. Because essay writing tests the student's ability to generate his own correct sentences--rather than respond to errors in someone else's--we think it to be potentially the most valid instrument. We recognize also, however, that even with marking essays with a checklist (as we did), certain subjective judgements are inevitable. This is less true when we look at mechanics rather than style, but such things as logic, organization, etcetera are all components of an essay's final grade. Thus once again we were faced with making a compromise in our test instruments; and again, given the various constraints of research, we believe we have been realistic.

We also broke down the essay grades one further degree by looking at the average mark for the last three essays. We reasoned that these essays reflected the student's skills after he had gone through all or nearly all of the lessons. Although we found no significant differences in both sets of essay scores, we did detect a slight trend in favour of the CAI with D&P group. Their improvement was the largest in absolute terms. Our research next year will try to investigate what factors led to this result.

We might point out again the effects of the strike. It was our original intention to use the PI and CAI lessons during the first part of the semester only; the skills taught then would be refined and reinforced by essay writing in the last part of the course. But we were forced to abandon this plan because of the time lost; students worked on the lessons virtually up to the day of the final test. Since we think that neither CAI nor PI is suited for more than an adjunct role in a total writing program, we regret that this material had to be relied upon so heavily.

Neither were we fully satisfied with the population of our experiment. It had been our intention to deal only with remedial students who were native speakers. About 28% of our subjects were actually second-language students whose problems were quite different than those for whom the lessons were designed. (In Appendix II, we include a discussion on the general pedagogic problems this raised.) The Chi-square test that we carried out did not reveal that any of the treatment groups had a disproportionately high number of second-language students, and thus we consider our results

valid. But we think that overall achievement would have been greater had we not had to face this unexpected variable. In passing, we recommend that the College reconsider its policy of placing second-language students in the regular Core stream.

That significant differences in achievement were not found is both puzzling and yet quite consistent with the research literature. The theoretical advantages of CAI--dynamic, individualizing, motivating and so forth--should have produced important changes in learning. But our study finds no real evidence of this; and one cannot say that similar other studies have been more successful. Does this suggest that CAI should be abandoned? After all, what do we gain from the time and money invested that other delivery systems can't give us? As designers and implementers of both CAI and PI material this year, we would like to speak to that question.

From the perspective of the teacher-designer who wishes to produce materials for his population, CAI does have a number of clear advantages. Like PI, a well-designed CAI lesson can teach students, and in so doing, free teachers to concentrate on more complex tasks. Like PI, CAI allows the student to work at his own rate on those areas where his needs are greatest. And like PI, CAI can give students opportunities for extra practice. But PI soon reaches a point where additional materials become very difficult to deliver. Revisions to a CAI program require simply accessing the lesson and entering a new line. The updated lesson is available to students without delay. To revise PI requires the production of new masters, new

printing, collation, and so forth. Given these conditions, one hesitates to make major revisions to PI (and minor corrections like typographical errors get even shorter shrift in practice). Our experience this year demonstrated that the extra time and money spent producting PI texts--even with the best and most diligent intentions--did not result in a truly satisfactory product. All other things being equal, then, CAI is far more convenient to update, revise, and expand.

Historically, cost has been the most potent argument against CAI's implementation. This is no longer convincing. Hardware costs have dropped dramatically and all indications are that they will fall still more. Our expenses this year largely were incurred in the design of the lessons, not the purchase or operation of equipment. Moreover, the interactive nature of CAI allows a course author to write his lessons directly at the terminal, minimizing the need for secretarial and other support help. We estimate the cost of producing the PI texts--exclusive of lesson design--to be about $1200. The cost breakdown goes as follows:

Typing from computer listing	15 hours
Producing thermofax from masters	1
Running off dittos	2
Collating	4
Total	20 hours

At $3.50 an hour, personnel costs come to $70 per lesson, to which must be added about $25 for materials, and between $2.50 and $5.00 for stapling. Thus each PI lesson cost the College $100.00, or $1200 in all.* Should we want to use the lessons again, we would still incur six hours of time per lesson for running off dittos and collating (or $21.00). The stapling and

*These are very conservative estimates, one should note.

paper charges would bring the price to about $50 per lesson reprint, or $600.00 for all twelve. In contrast, the CAI groups required about $20 worth of paper (2000 sheets at about $10/1000), and the use of a CRT would reduce this to nothing. Line charges for the terminal cost the College about $250. Whether or not to include CPU time is a question. On one hand, the Ministery of Education provides computer services to all CEGEP's without charge. It seems to us wasteful not to use these facilities. It is true, of course, that the CPU costs are borne ultimately by the taxpayer, if not by the College directly. But it simply does not follow that given the conditions under which the College works, costs for using CAI are prohibitive. The major expense, as we have said, is the design of the actual courseware, but such costs can be amortized by extending the use of CAI lessons to many more students. On a per-student basis, we think such instruction will prove to be cost effective.

A third major advantage CAI has over PI (and indeed any other medium we know of) is its ability to gather data on how individuals learn and how, therefore, we should teach. The encoding of a CAI lesson presupposes an exhaustive preparation. Although careful planning is the hallmark of all good instruction, only CAI, we think, makes it a necessary condition. Our CAI lessons were designed so as to record each student response in terms of a 1 or 0 (for right or wrong). We think this data will be invaluable to us when we revise next year. Defective questions will be eliminated and successful ones will act as models for additions. Moreover, we hope to go back over this data and try to find out what we

can about the learning styles of our population. As more students use the lessons, more data will be generated, and we hope to refine our procedures to make such information even more relevant to the problems of how and why students acquire--or fail to--the skills we wish to teach. PI simply does not lend itself to such research. The computer operates unobtrusively; students have no sense that their responses are being recorded. To ask the PI student to return his book to the teacher so that his answers can be charted--a nightmare simply in terms of logistics--would very likely change the PI experience substantially. We think that such data should be gathered and that CAI has a clear advantage in doing so. In fact, we think that such data is the key to solving the basic problems associated with CAI research. At first blush, comparisons between CAI and PI seem quite fruitful--two different delivery modes for similar populations, using essentially parallel materials appears to mean that only one independent variable is being manipulated. But, in fact, we have come to recognize that there are a great many more variables at work than just the delivery system. CAI doesn't merely branch where PI doesn't (or merely branch more rapidly than a scrambled PI text): it interacts dynamically with the student, and that dynamic interaction provides an experience different in kind that PI, rather than just degree. As each student traces his own path through a CAI program--with as many possible combinations as there are learners and frames--the experience is far, far different from a linear progression. How that difference makes itself manifest, what can we say about why and how these paths evolve, what the implications are for different sorts of branching--

these are the issues that now seem to us to be the most pressing. We believe we should be less concerned with whether or not CAI is better than other systems; we should rather ask the questions about what elements of CAI work for what materials, for what populations, and under what conditions. Indeed how can the computer's capacity to store and retrieve information literally in nanoseconds be exploited? How can a computer's potential for artificial intelligence be used in CAI? What sort of learning theory should we call upon when writing CAI programs that is different from that used in PI or other modes of instruction? What sort of data should we be collecting about the people who go through a program--simply overall achievement? The number of errors made? The places where these errors occur? What sorts of remedial strategies can be matched to particular styles of learning? Research must begin to address itself to these kinds of questions rather than on what may well be a somewhat spurious issue of what is better, PI or CAI. The truth, we suspect, is not a matter of better or worse, but of essential differences that would lead one to employ one medium under one set of circumstances and not another.

Perhaps the reason the literature is so inconclusive is that researchers have tended to concentrate more on the differences between the two treatments than on differences within each treatment. This is the sort of research problem that is neither new nor peculiar to comparisons between CAI and PI. As early as 1928, R.B. Spence wrote:

> The problem that must be solved is not the question "Is Method A better than Method B or Method C?" but rather, "What are the conditions under which Method

> A produces more effective results? What are the situations where Method B is best, etc."
> (Spence, 1928, p.426)

Almost forty years later, Briggs et al (1967) argued that the lack of significant differences between media in making learning effective might well be because "each of the media compared was more effective for some elements of instruction and less effective for others" (p.24), the net result often being that each one canceled out the other's benefits. Chu and Schramm (1967, p.180), reporting on educational television, concluded that it was not the"fanciness that might be peculiar to television" that made for effectiveness, but rather "good teaching . . . qualities like simplicity, good organization, motivation . . .". Chu and Schramm's definition of good teaching may not coincide with everyone's nor be specific enough; it does, however, underline the point that without sound, well-articulated, and well-researched pedagogical principles informing instruction, the choice of particular media becomes less critical. The conclusion reached by Chu and Schramm's (1967) study seems to us to be especially to the point here: given favourable conditions pupils can learn from any instructional medium now available. The refinement of the meaning of "favourable conditions" should be the first order of research if we are to come to understand what is truly effective in PI and CAI. And as we've said before, CAI, requiring the course author to articulate his strategies fully before encoding, is especially useful in helping us towards that difficult notion of "favourable conditions". W.H. Allen (1971) also is worth quoting:

> The study of three-way interaction of stimulus, task, and learner is extremely complex . . . the folly of assigning generalized and all-inclusive attributes to specific classes of media under all conditions is finally being appreciated . . . [proper research will lead] to the evolution of taxonomies of unique media effects so that we can predict that the use of a particular instructional medium will lead to specified learning outcomes with different kinds of learners.

We think now that future work must pay more attention to what constitutes "good teaching" for CAI or PI. We should be less concerned with demonstrating the superiority of one mode over the other, but instead try to discriminate between the elements of each that makes it uniquely effective.

In addition to these general research questions, we would also like to recommend a number of changes for implementing the material. We think sharper diagnostic tools are needed before the student begins to work his way through the lessons. Our experimental design this year did not use a pretest because we were interested in the relative effects of the different media rather than absolute gains. We therefore decided not to risk sensitization by the pretest. In a classroom context, however, diagnostic pretesting would enable the teacher to identify the particular problems of the students and direct them accordingly. (If a reliable cognitive style mapping system were available, this too should be incorporated into the pretesting procedure.) We also think that the lessons should be reinforced by class time; again, our experimental design did not admit this extra variable. But we do think that students would benefit by such extra help. We also suggest that an additional test on each lesson be administered,

with the scores to count towards the final grade. The criterion tests which are part of the lessons are really intended as vehicles to inform both student and teacher about the progress being made; they were not designed as tools of formal evaluation. To change them in this way would make the lessons another obstacle to be hurdled, rather than learning for its own sake. We do not wish to press this last point; how the material will be used must finally be left to the professional judgement of the classroom teacher.

Indeed, we wish to emphasize that we have never considered the lessons to be anything more than an adjunct activity. They are certainly no substitute for constant and personal supervision of writing. At best, CAI or PI can make a beginning on the serious and difficult task of teaching writing correctness. To claim that twelve or so hours at a computer terminal--or, for that matter, fifteen weeks in a classroom--can produce an accomplished writer is simply to mislead. We were not always able to produce even merely competent writers, an experience familiar to teachers of composition. Again and again, we experienced the frustration of all teachers whose students fail to make that quantum leap to perfection. We would like to report that God is in his heaven, and all's right with the composition world; alas, we cannot. But we do think we have made an important start in teaching the crucial skill of writing. We think the computer has a role to play in collegial education, and we remain committed to an objective exploration of what that role can be.

APPENDIX I

STUDENT EVALUATION OF TERMINAL AND BOOKLET GRAMMAR LESSONS

Please answer the following questions about the lessons you did this semester in Effective Writing and Linguistic Structures. Your responses will provide valuable information for evaluating and changing them. Thank you very much for your help.

Fill in the appropriate answers on the sheet provided.

1) I took these lessons in the following format:

C	XC	L	XB	B
1	2	3	4	5

2) My first language is:

English	French	Other
1	2	3

3) The language in which I did my previous schooling was:

English	French	Other
1	2	3

Use the following key for the rest of the questions:

Strongly agree	Agree	Uncertain	Disagree	Strongly Disagree
1	2	3	4	5

When using the lessons, I generally experienced:

4) fun
5) frustration
6) a challenge
7) annoyance
8) involvement
9) confusion
10) boredom
11) encouragement
12) control
13) success

14) Using the lessons was an enjoyable experience.

15) The method of presenting the material in the lessons distracted me from learning.

16) The major points of the lessons were made clear.

17) I learned what the lessons tried to teach.

18) I already knew the material covered in the lessons.

19) I think these lessons helped me to learn the material more thoroughly than any other form of instruction would have.

20) These lessons helped me to learn more efficiently than other ways that I have been taught.

21) The lessons were more interesting than traditional classroom/textbook instruction.

22) I spent too much time on the lessons.

23) I would have preferred spending more time on the material in class.

24) I would like to take another course using this same approach.

25) The pace at which the lessons came was:

Much too fast	Too fast	About right	Too slow	Much too slow
1	2	3	4	5

26) The content of the lessons was:

Much too difficult	Too difficult	About right	Too easy	Much too easy
1	2	3	4	5

27) Write any other comments in the space below and return this sheet:

APPENDIX II

SECOND-LANGUAGE STUDENTS

Any writing program designed to teach the mechanics of writing has to be based on a very clear set of objectives defined in advance--primarily, what has to be taught and what method should be used to teach it. The experience at this college is that students enter with widely-varying writing (and reading) experience and with an enormous and somewhat unpredictable range of writing problems. We cannot assume that errors common to native speakers will be shared by second-language students. The non-native speaker may have very fundamental problems with English idiom (the preposition and articles, for example) which no native speaker would manifest in his writing.

A further complication in the case of preparing a writing program tailor-made to the needs of our students at Snowdon has been the wide range of backgrounds encountered by the teacher in the classroom. The following represents the range of mother languages in a class of 24 in Linguistic Structures (603-404-75) this past term:

Chinese	17
Vietnamese	3
Italian	1
French	1
Iranian	1
Turkish	1

Of these, some had been educated at least partly in Montreal high schools, while others were recently-arrived immigrants. Again, some of the students

had already had one, two, or even three previous remedial-level courses (Effective Writing or Linguistic Structures) while others were taking their first college-level English course. This heterogeneity is a feature of many remedial courses, but is particularly marked in Linguistic Structures. The levels of assignment, complexity of topics in English grammar and usage, and the grading standards that can reasonably apply to such a heterogeneous class all had to be determined and monitored by the teacher.

Ideally, if the College can predict admission of a significant number of students from one language background and in need of remedial work in English, a more specialized set of lessons suited to them should be designed. Again, specialists in second-language teaching would have to be employed to serve these students. It ought to be noted that in the other remedial course, Effective Writing, there also may be as many as one-half of the students whose main problems arise from linguistic confusion with a different language spoken in the home. Surveying three different Effective Writing courses taught by one teacher in the last three semesters, the following chart is instructive:

#1 (H77)		#2 (A77)		#3 (H78)	
English	5	English	9	English	11
Italian	4	Italian	5	Italian	2
French	2	French	6	French	8
Chinese	4			Chinese	2
Total	15	Total	20	Total	23

These figures represent only an approximation, as many students are effectively bilingual or trilingual apart from idiomatic or usage problems. But

they do indicate that at this college, increasingly, remedial composition courses are complicated by the special problems of second-language teaching.

REFERENCES

ABBOUD, Victorine
BUNDERSON, Victor C.

A Computer Instruction Program in the Arabic Writing System. Austin, Tex.: The University of Texas. Austin Computer-Assisted Instruction Laboratory, 1971.- (Technical Report, No. 4).

ALLEN, W.H.

"Instructional Media Research: Past, Present and Future." AV Communitcations Review, vol. 19, 1971.

BLITZ, Allan N.

An Investigation of the Ways in Which Personality Characteristics Affect Performance on Computer-Assisted Instruction and Programmed Text. Ed. D. thesis. University of Kentucky, 1972, 128 p.

BOSSONE, Richard M.
WEINER, Max

Three Modes of Teaching Remedial English: A Comparative Analysis; a Pilot Study. New York: City University of New York. Bernard Baruch College and Graduate School, 1973, 87 p.
ERIC ED 74514

BRIGGS, L.G. et al

Instructional Media: A Procedure for the Design of Multimedia Instruction, A Critical Review of Research, and Suggestions for Future Research. Pittsburgh: American Institute for Research, 1967.

BROWN, B.

"Effect of CAI and PI on Attitude." Contemporary Education, vol. 40, no. 5, Terre Haute: Indiana State University, School of Education, 1969.

CALDWELL, Robert McKinley

A Comparison of a Programmed Text and a Computer Based Display Unit to Teach Reading Skills to Semi-Literate Adolescents. Pennsylvania State University, 1973, 219 p.

CALLIS, Robert
JOHNSON, Willouby

Missouri College English Test, New York, 1965.

CHU, G.C.
SCHRAMM, W.

Learning from Television: What the Research Says. Washington, D.C.: National Association of Broadcasters, 1967.

COULSON, J.E. et al

"Effects of Branching in a Computer Controlled Auto-Instructional Device." Technical Memorandum. Santa Monica, Calif.: Systems Development Corporation, 1961, 8 p.

CROWDER, Norman A.

"On the Differences between Linear and Intrinsic Programming" in John P. De Cecco, ed. Educational Technology, Readings in Programmed Instruction. New York: Holt Rinehart and Winston, 1964, pp. 142-152.

DEARBORN, R.O.
CALIGAERT, P.

The Value of Hardcopy Output in Computer-Administered Programmed Programmed Instruction. Programmed Learning and Educational Technology, vol. 9, 1972.

DICK, Robert et al

Personnel Technology: Using an Inexpensive Computer-Based System to teach Performances Oriented Skill. Goleta, Calif.: Human Factors Research, Inc., 1974, 166 p. - (Technical Report no. 1728-1).

DICK, Walter
LATTA, Raymond

Comparative Effects of Ability and Presentation Mode in Computer-Assisted Instruction and Programmed Instruction. Audio Visual Communication Review, vol. 18, no. 1, Spring 1970, pp. 33-45.

GRUBB, Ralph E.
SELFRIDGE, Lenore D.

"Computer Tutoring in Statistics." Computer and Automation, vol. 13, no. 3, March 1964, pp. 20-26.

HAUGH, Oscar M.
BROWN, James I.

College English Placement Test, Boston, 1969.

HOMEYER, F.C.

"Development and Evaluation of an Automated Assembly Language Teacher." Technical Report no. 3, Computer-Assisted Instruction Laboratory, University of Texas, 1970.

HONG, Sunion Theordore

An Empirical Study of the Effectiveness of Programmed Instruction and Computer-Assisted Instruction in Elementary Accounting. Ph.D. thesis. New York: New York University. Graduate School of Business Administration, 1972, 161 p.

JOHNSON Warner Goodwin

The Effects of Computer-Assisted Instruction and Programmed Instruction on the Achievement and Attitude of Ninth Grade General. Philadelphia: Temple University, 1973, 242 p.

LENN, Peter D.
MASER, Thomas F.

Computer Assisted/Programmed Instruction on the Law. Final Report. San Francisco: American Analysis Corp., 1971, 51 p. ERIC ED 58724

McMURRY, James G.

An Experimental Comparison of Self-Instructional Materials Versus Traditional Methods in Teaching Related Information in the Comprehensive General Shop. Unpublished Doctor's Dissertation, University of Missouri, 1964.

MULL, Patricia Ann

A Comparison of Computer-Assisted Instruction and a Programmed Workbook to Teach the Initial Teaching Alphabet and Trans-Literation to Inservice Teachers. Ph.D. Thesis. University Park: The Pennsylvania State University Press, 1973, 145 p.

ROID, Gale H.

Covariates of Learning in Computer-Assisted Instruction. Paper presented at the Annual Meeting of the American Educational Research Association, New York, February 1971. Montreal: McGill University, 1971, 8 p.

SCHOLER, Marc

Recherche en Enseignement Programme Livresque et Informatise (Operation REPLI), Rapport Final I, Montreal 1976.

SKINNER, B.F.

Teaching Machines in A.A. Lumsdaine and R. Glaser, eds. Teaching Machines and Programmed Learning: A Source Book. Washington, D.C.: National Education Association, Department of Audio-visual Instruction, 1960, pp. 137-158.

SPENCE, R.B.

"Lecture and Class Discussion in Teaching Educational Psychology." Journal of Educational Psychology, 1928, 19 p.